A Class of Semidefinite Programs with a rank-one solution

Guillaume Sagnol
Inria Saclay – Île-de-France & Centre de Mathématiques Appliquées (CMAP),
École Polytechnique, France.
guillaume.sagnol@inria.fr

July 9, 2010

Abstract

We show that a class of semidefinite programs (SDP) admits a solution that is a positive semidefinite matrix of rank at most r, where r is the rank of the matrix involved in the objective function of the SDP. The optimization problems of this class are semidefinite packing problems, which are the SDP analogs to vector packing problems. Of particular interest is the case in which our result guarantees the existence of a solution of rank one: we show that the computation of this solution actually reduces to a Second Order Cone Program (SOCP). We point out an application in statistics, in the optimal design of experiments.

Keywords SDP, Semidefinite Packing Problem, rank 1-solution, Low-rank solutions, SOCP, Optimal Experimental Design, Multiresponse experiments.

1 Introduction

In this paper, we study *semidefinite packing problems*. The latter, which are the semidefinite programming (SDP) analogs to the packing problems in linear programming, can be written as:

$$\begin{aligned} \max \quad & \langle C, X\rangle \\ \text{s.t.} \quad & \langle M_i, X\rangle \leq b_i, \qquad i \in [l], \\ & X \succeq 0, \end{aligned} \tag{1}$$

where $C \succeq 0$, and $M_i \succeq 0,\ i \in [l]$. The notation $X \succeq 0$ indicates that X belongs to the set $\mathbb{S}_n^+$ of $n \times n$ positive semidefinite matrices. Similarly, $X \succ 0$ stands for $X \in \mathbb{S}_n^{++}$, the set of $n \times n$ positive definite matrices. The space of $n \times n$ symmetric matrices $\mathbb{S}_n$ is equipped with the inner product $\langle A, B\rangle = \mathrm{trace}(A^T B)$. We also make use of the standard notation $[l] := \{1, \ldots, l\}$, and we use boldface letters to denote vectors.

Semidefinite packing problems were introduced by Iyengar, Phillips and Stein [8]. They showed that these arise in many applications such as relaxations of combinatorial optimization problems or maximum variance unfolding, and gave an algorithm to compute approximate solutions, which is faster than the commonly used interior point methods.

Our main result is that when the matrix C is of rank r, Problem (1) has a solution that is of rank at most r (Theorem 2). In particular, when $r = 1$, the optimal SDP variable X can be factorized as $\boldsymbol{x}\boldsymbol{x}^{\boldsymbol{T}}$, and we show that finding $\boldsymbol{x}$ reduces to a Second-Order Cone Program (SOCP) which is computationally more tractable than the initial SDP. We present this result and some applications in Section 2. Then, we extend our result to a wider class of semidefinite programs (Theorem 4), in which not all the constraints are of packing type. The proofs of the theorems are given in Section 4.

2 Main result and consequences

In this section, we state the main result of this article and point out an application to statistics. We also discuss the significance of our result for combinatorial optimization problems (the hypothesis on the rank of the matrix C appears to be very restrictive). The two theorems of this section are proved in Section 4.

2.1 The main result

We start with an algebraic characterization of the semidefinite packing problems that are feasible and bounded.

Theorem 1. *Problem* (1) *is feasible if and only if every b_i is nonnegative. Moreover if Problem* (1) *is feasible, then this problem is bounded if and only if the range of C is included in the range of $\sum_i M_i$.*

The main result of this article follows:

Theorem 2. *We assume that the conditions of Theorem 1 are fulfilled, so that Problem* (1) *is feasible and bounded. If* $\operatorname{rank} C = r$*, then the semidefinite packing problem* (1) *has a solution that is a matrix of rank at most r. In addition, if $\min_{i \in [l]} \operatorname{rank} M_i = \overline{r}$, then every solution of Problem* (1) *must be of rank at most $n - \overline{r} + r$.*

A consequence of our theorem is that when the matrix in the objective function is of rank 1 ($C = cc^T$), the computation of the solution X of Problem (1) reduces to the computation of a vector $\boldsymbol{x}$ such that $X = \boldsymbol{x}\boldsymbol{x}^{\boldsymbol{T}}$. The next result shows that this can be done very efficiently by a Second Order Cone Program (SOCP).

Corollary 3. *We assume that the conditions of Theorem 1 are fulfilled, and that $C = \boldsymbol{c}\boldsymbol{c}^{\boldsymbol{T}}$ for a vector $\boldsymbol{c} \in \mathbb{R}^n$ (i.e.* $\operatorname{rank}(C) = 1$*). Then, Problem* (1) *reduces to the SOCP:*

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & \boldsymbol{c}^{\boldsymbol{T}}\boldsymbol{x} \\ s.t. \quad & \|A_i \boldsymbol{x}\|_2 \leq \sqrt{b_i}, \qquad i = 1 \in [l], \end{aligned} \tag{2}$$

where the matrices A_i are such that $M_i = A_i^T A_i$. Moreover, if $\boldsymbol{x}$ is any optimal solution of Problem (2)*, then $X = \boldsymbol{x}\boldsymbol{x}^{\boldsymbol{T}}$ is an optimal solution of Problem* (1)*, and the optimal value of* (1) *is $(\boldsymbol{c}^{\boldsymbol{T}}\boldsymbol{x})^2$.*

Proof. The SOCP (2) is simply obtained from (1) by substituting $\boldsymbol{x}\boldsymbol{x}^{\boldsymbol{T}}$ to X and $A_i^T A_i$ to M_i. The objective function $\langle C, X\rangle$ becomes $(\boldsymbol{c}^{\boldsymbol{T}}\boldsymbol{x})^2$, and we can remove the square by noticing that $\boldsymbol{c}^{\boldsymbol{T}}\boldsymbol{x} \geq 0$ without loss of generality, since if $\boldsymbol{x}$ is optimal, so is $-\boldsymbol{x}$. □

Remark 1. The ratio between the optimal value of Problem (1) and the value of its best solution of rank one has been studied By Nemirovski, Roos, and Terlaky [11]. They show that the value v^* of the SDP and the value v_1^* of its best rank-one solution satisfy:

$$v^* \geq v_1^* \geq \frac{1}{2\ln(2l\mu)} v^*, \qquad \text{where } \mu = \min(l, \max_{i \in [l]} \operatorname{rank} M_i). \tag{3}$$

This ratio can be considerably reduced in particular configurations, but to the best of our knowledge, the fact that the gap in (3) vanishes when the matrix C in the objective function is of rank 1 is new, except in the particular case in which every M_i is of rank 1, too [15].

The present work grew out from an application to networks [4], in which the traffic between any two pairs of nodes must be inferred from a set of measurements. This can be modeled by the theory of optimal experimental design, which leads to a large SDP. Standard solvers relying on interior points methods, like SeDuMi [18], cannot handle problems of this size. However, in a followup work relying on the present reduction to an SOCP [17], we solve within seconds the same instances in SeDuMi. We next present this application to statistics.

2.2 Application to the optimal design of experiments

An interesting application arises in statistics, in the design of optimal experiments (for more details on the subject, the reader is referred to Pukelsheim [14]). An experimenter wishes to estimate the quantity $\boldsymbol{c^T\theta}$, where $\boldsymbol{\theta}$ is an unknown $n-$dimensional parameter, and $\boldsymbol{c}$ is a vector of n coefficients. To this end, she disposes of l available experiments, each one giving a linear measurement of the parameter $\boldsymbol{y_i} = A_i\boldsymbol{\theta}$, up to a (centered) measurement noise. If the amount of experimental effort spent on the i^{th} experiment is w_i, it is known that the variance of the best linear unbiased estimator for $\boldsymbol{c^T\theta}$ is $\boldsymbol{c^T}(\sum_i w_i M_i)^\dagger \boldsymbol{c}$, where $M_i = A_i^T A_i$, and $M^\dagger$ denotes the Moore-Penrose inverse of M. The problem of distributing the experimental effort so as to minimize this variance is called the "$\boldsymbol{c}-$optimal problem", and can be formulated as:

$$\begin{aligned} \min_{\boldsymbol{w}\geq \boldsymbol{0}} \quad & \boldsymbol{c^T}(\sum_i w_i M_i)^\dagger \boldsymbol{c} \\ \text{s.t.} \quad & \sum_{i=1}^{l} w_i = 1. \end{aligned} \tag{4}$$

It is classical to reformulate this problem as a semidefinite program, by using the Schur complement lemma and duality theory (see e.g. [15, 16]). The $c-$optimal SDP already appeared in Pukelsheim and Titterington [13], hidden under a more general form:

$$\begin{aligned} \max \quad & \boldsymbol{c^T} X \boldsymbol{c} \\ \text{s.t.} \quad & \langle M_i, X\rangle \leq 1, \qquad i \in [l], \\ & X \succeq 0. \end{aligned} \tag{5}$$

In this problem, the design variable $\boldsymbol{w}$ is proportional to the dual variable associated to the constraints $\langle M_i, X\rangle \leq 1$. Note that this is a semidefinite packing problem, in which the matrix defining the objective function has rank 1 ($C = \boldsymbol{cc^T}$). More generally, if we want to estimate simultaneously r linear functions of the parameter $\boldsymbol{\zeta} = (\boldsymbol{c_1^T\theta}, \ldots, \boldsymbol{c_r^T\theta})$, the best unbiased estimator is now a $r-$dimensional vector with covariance matrix

$$\operatorname{Cov}_{\boldsymbol{w}}(\hat{\boldsymbol{\zeta}}) := K^T(\sum_{k=1}^{l} w_k M_k)^\dagger K,$$

where $K = [\boldsymbol{c_1}, \ldots, \boldsymbol{c_r}]$. Several criteria can be used for this experimental design problem. Popular ones are the $A-$criterion and the $E-$criterion, which aim at minimizing respectively the trace and the largest eigenvalue of $\operatorname{Cov}_{\boldsymbol{w}}(\hat{\boldsymbol{\zeta}})$. These optimization problems can also be formulated as semidefinite packing problems. For $A-$optimality, this *packing* formulation is given in [16]:

$$\begin{aligned} \max \quad & \boldsymbol{\tilde{c}^T} X \boldsymbol{\tilde{c}} \\ \text{s.t.} \quad & \langle \tilde{M}_i, X\rangle \leq 1, \qquad i \in [l], \\ & X \succeq 0, \end{aligned} \tag{6}$$

where $\tilde{\boldsymbol{c}} = [\boldsymbol{c_1^T}, \ldots, \boldsymbol{c_r^T}]^T$, and $\tilde{M}_i$ is a block-diagonal matrix which contains r times the block M_i on its main diagonal. The matrix in the objective function is of rank 1 ($C = \tilde{\boldsymbol{c}}\tilde{\boldsymbol{c}}^{\boldsymbol{T}}$), and so Problem (6) reduces to a SOCP by Corollary 3. This reduction is of great interest for the computation of optimal experimental designs, because SOCP solvers are much more efficient than SDP solvers, and take advantage of the sparsity of the matrices A_i (whereas the matrices $M_i = A_i^T A_i$ used in the original SDP formulation (6) are *not very sparse* in general).

The $E-$optimal design SDP is presented in [20] (for the special case in which $C = I$), and takes exactly the form of the semidefinite packing problem (1), with $b_i = 1$ for all $i \in [l]$ and $C = KK^T = \sum_{i=1}^r \boldsymbol{c_i}\boldsymbol{c_i}^T$. Here, the matrix C has rank r, and so Theorem 2 indicates that the $E-$optimal design SDP has a solution that is a matrix of rank at most r. This suggests the use of specialized low rank solvers for this SDP when r is small (cf. Remark 2), which can lead to a considerable improvement in terms of computation time.

Remark 2. We point out that solutions of small rank of semidefinite programs have been extensively studied over the past years. Barvinok [3] and Pataki [12] discovered independently that any SDP with l constraints has a solution X^* whose rank is at most

$$r^* = \left\lfloor \frac{\sqrt{8l+1}-1}{2} \right\rfloor,$$

where $\lfloor \cdot \rfloor$ denotes the integer part. This was one of the motivations of Burer and Monteiro for developing the SDPLR solver [5], which searches a solution of the SDP in the form $X = RR^T$, where R is a $n \times r^*$ matrix. The resulting problem is non-convex, and so the augmented Lagrangian algorithm proposed in [5] is not guaranteed to converge to a global optimum. However, it performs remarkably well in practice, and some conditions which ensure that the returned solution is an optimum of the SDP are provided in [6]. Our result shows that for a semidefinite packing problem in which the matrix C has rank r, one can force the matrix R to be of size $n \times r$ (rather than $n \times r^*$), which can lead to considerable gains in computation time when r is small.

2.3 Relation with combinatorial optimization

SDP relaxations of combinatorial optimization problems have motivated the authors of [8] to study semidefinite packing problems. Hence, we discuss the significance of our result for this class of problems in this section.

Semidefinite programs have been used extensively to formulate relaxations of NP-hard combinatorial optimization problems after the work of Goemans and Williamson on the approximability of MAXCUT [7]. These SDP relaxations often lead to optimal solutions of the related combinatorial optimization problems whenever the solution of the SDP is of small rank. As shown by Iyengar et. al. [8], SDP relaxations of many combinatorial optimization problems can be cast as semidefinite packing programs. Our result therefore identifies a subclass of combinatorial optimization problems which are solvable in polynomial time. Unfortunately, this promising statement only helped us to identify trivial instances so far. For example, the MAXCUT semidefinite packing problem [8] yields an exact solution of the combinatorial problem whenever it has a rank 1 solution. The matrix C in the objective function of this SDP is the Laplacian of the graph, and so it is known that

$$\operatorname{rank} C = N - \kappa,$$

where N is the number of vertices and κ is the number of connected components in the graph. Our result therefore states that if a graph of N vertices has $N-1$ connected components, then it defines a MAXCUT instance that is solvable in polynomial time. Such graphs actually consist in a pair of connected vertices, plus $N-2$ isolated vertices, and the related MAXCUT instance is trivial.

Another limitation for the application of our theorem in this field is that most semidefinite packing problems arising in combinatorial optimization (including but not limited to the Lovász ϑ function

SDP [10] and the related Szegedy number SDP [19], the vector colouring SDP [9], the sparsest cut SDP [1] and the sparse principal components analysis SDP [2]) can be written in the form of (1), with an additional trace equality constraint $\text{trace}(X) = 1$. In fact, we can show that if such an "equality constrained" problem is strictly feasible, then it is equivalent to the following "classical" semidefinite packing problem:

$$\begin{aligned} \max \quad & \langle C + \lambda I, X\rangle - \lambda \\ \text{s.t.} \quad & \langle M_i, X\rangle \leq b_i, \qquad i \in [l], \\ & \text{trace} X \leq 1, \\ & X \succeq 0, \end{aligned} \tag{7}$$

where λ is any scalar larger than $|\lambda^*|$, where λ^* is the optimal Lagrange multiplier associated to the constraint $\text{trace}(X) = 1$ (we omit the proof of this statement which is of secondary importance in this article). Since $C + \lambda I$ is a full rank matrix, our result does not seem to yield any valuable information for this class of problems.

3 Extension to "combined" problems

The proof of our main result also applies to a wider class of semidefinite programs, which can be written as:

$$\begin{aligned} \max_{X,Y,\boldsymbol{\lambda}} \quad & \langle C, X\rangle + \langle R_0, Y\rangle + \boldsymbol{h_0}^T \boldsymbol{\lambda} \\ \text{s.t.} \quad & \langle M_i, X\rangle \leq b_i + \langle R_i, Y\rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda}, \qquad i \in [l], \\ & X \in \mathbb{S}_n^+,\ Y \in \mathbb{S}_p^+,\ \boldsymbol{\lambda} \in \mathbb{R}^q, \end{aligned} \tag{8}$$

where **every matrix M_i and C are positive semidefinite, while the R_i are *arbitrary symmetric* matrices**. The vectors $\boldsymbol{h_i}$ are in $\mathbb{R}^q$. We denote by H the $q \times l$ matrix formed by the columns $\boldsymbol{h_1}, \ldots, \boldsymbol{h_l}$. The dual of Problem (9) is:

$$\begin{aligned} \min_{\boldsymbol{\mu} \geq 0} \quad & \boldsymbol{b}^T \boldsymbol{\mu} \\ \text{s.t.} \quad & \sum_{i=1}^{l} \mu_i M_i \succeq C, \\ & R_0 + \sum_{i=1}^{l} \mu_i R_i \preceq 0. \\ & \boldsymbol{h_0} + H\boldsymbol{\mu} = \boldsymbol{0}. \end{aligned} \tag{9}$$

In the next theorem, we show that if the dual problem is strictly feasible, and the primal is feasible, then a solution of rank at most r exists. This theorem is proved in Section 4.

Theorem 4. *We assume Problem* (9) *is strictly feasible and Problem* (8) *is feasible, i.e. the following two conditions hold:*

(i) $\exists \overline{\boldsymbol{\mu}} \geq \boldsymbol{0} : \ \sum_i \overline{\mu}_i M_i \succ C,\ R_0 + \sum_i \overline{\mu}_i R_i \prec 0,\ \boldsymbol{h_0} + H\overline{\boldsymbol{\mu}} = \boldsymbol{0},$

(ii) $\exists (\overline{Y}, \overline{\lambda}) \in \mathbb{S}_p^+ \times \mathbb{R}^q : \ \forall i \in [l],\ b_i + \langle R_i, \overline{Y}\rangle + \boldsymbol{h_i}^T \overline{\boldsymbol{\lambda}} \geq 0.$

Then, Problem (8) *has a solution* $(X, Y, \boldsymbol{\lambda})$ *in which*

$$\text{rank } X \leq r := \text{rank } C.$$

In addition, if $\min_{i\in[l]} \text{rank } M_i = \overline{r}$*, then every solution* $(X, Y, \boldsymbol{\lambda})$ *of Problem* (8) *must be such that* $\text{rank } X \leq n - \overline{r} + r$.

Remark 3. In the SDP (8), we have introduced an unbounded vector variable $\boldsymbol{\lambda}$. The reader may wonder why, since we could have expressed $\boldsymbol{\lambda}$ as the diagonal of the difference of two positive semidefinite matrices instead. Although subtle, introducing an unbounded variable *does* make a difference, since otherwise the dual strict feasibility assumption (i) would require $\boldsymbol{h_0} + H\overline{\boldsymbol{\mu}} < \boldsymbol{0}$ and $\boldsymbol{h_0} + H\overline{\boldsymbol{\mu}} > \boldsymbol{0}$ to hold at the same time (which of course is not possible).

As in the previous section, we have a result of reduction to a SOCP, which holds when C is of rank 1, every $R_i = 0$ and $\boldsymbol{h_0} = \boldsymbol{0}$.

Corollary 5. *Consider the following "combined" semidefinite packing problem:*

$$\begin{aligned} \max_{X\in\mathbb{S}_m,\ \boldsymbol{y}\in\mathbb{R}^{m'}} \quad & \langle C, X\rangle && (10)\\ s.t. \quad & \langle M_i, X\rangle \leq \boldsymbol{h_i}^T\boldsymbol{\lambda} + b_i, \qquad i \in [l],\\ & X \succeq 0. \end{aligned}$$

Assume that $C = \boldsymbol{cc^T}$ *has rank* 1*, and*

(i) $\exists\overline{\boldsymbol{\mu}} \geq \boldsymbol{0}: \ \sum_{i=1}^{l} \overline{\mu}_i M_i \succ C,\ H\overline{\boldsymbol{\mu}} = \boldsymbol{0}$,

(ii) $\exists\overline{\boldsymbol{\lambda}} \geq \boldsymbol{0}: \ H^T\overline{\boldsymbol{\lambda}} + \boldsymbol{b} \geq \boldsymbol{0}$,

where $H = [\boldsymbol{h_1}, \ldots, \boldsymbol{h_l}]$ *is the matrix whose columns are the vectors* $\boldsymbol{h_i}$*. Then, Problem* (10) *reduces to the following SOCP:*

$$\begin{aligned} \max_{\boldsymbol{x}\in\mathbb{R}^{m},\ \boldsymbol{y}\in\mathbb{R}^{m'}} \quad & \boldsymbol{c}^T\boldsymbol{x} && (11)\\ s.t. \quad & \left\| \begin{bmatrix} 2A_i\boldsymbol{x} \\ \boldsymbol{h_i}^T\boldsymbol{\lambda} + b_i - 1 \end{bmatrix} \right\|_2 \leq \boldsymbol{h_i}^T\boldsymbol{\lambda} + b_i + 1, \qquad i \in [l], \end{aligned}$$

where the matrices A_i *are such that* $M_i = A_i^T A_i$*. Moreover, if* $(\boldsymbol{x}, \boldsymbol{\lambda})$ *is a solution of Problem* (11)*, then* $(\boldsymbol{xx^T}, \boldsymbol{\lambda})$ *is a solution of Problem* (10)*, and the optimal value of* (10) *is* $(\boldsymbol{c}^T\boldsymbol{x})^2$.

Proof. Theorem 4 guarantees that Problem (10) has a solution $(X, \boldsymbol{\lambda})$ in which X has rank 1, i.e. $X = \boldsymbol{x}\boldsymbol{x}^T$ for some vector $\boldsymbol{x}$. The SOCP (11) is simply obtained from (8) by substituting $\boldsymbol{xx^T}$ to X and $A_i^T A_i$ to M_i. We also used the fact that for any vector $\boldsymbol{z}$ and for any scalar α, the hyperbolic constraint

$$\|\boldsymbol{z}\|_2^2 \leq \alpha$$

is equivalent to the second order cone constraint

$$\left\| \begin{bmatrix} 2\boldsymbol{z} \\ \alpha - 1 \end{bmatrix} \right\|_2 \leq \alpha + 1.$$

$\square$

Application: $\boldsymbol{c}-$optimal design of experiments with multiple resource constraints
In a more general setting than the classical $\boldsymbol{c}-$optimal design problem (4) presented in the previous section, $\boldsymbol{w}$ no longer represents the percentage of experimental effort to spend on each experiment, but describes some resource allocation to the available experiments, that is subject to multiple linear constraints $P\boldsymbol{w} \leq \boldsymbol{d}$, where P is a $q \times l$ matrix with nonnegative entries and $\boldsymbol{d}$ is a $q \times 1$ vector. This problem arises for example in a network-wide optimal sampling problem [17], where $\boldsymbol{w}$ is the vector of the sampling rates of the monitoring devices on all links of the network, and is subject to linear constraints that limit the overhead of the routers. We will show that this problem is a "combined" semidefinite packing problem which reduces to an SOCP. The resource constrained $\boldsymbol{c}-$optimal design problem reads as follows:

$$\begin{aligned} \min_{\boldsymbol{w}\geq\boldsymbol{0}} \quad & \boldsymbol{c}^{\boldsymbol{T}}(\sum_i w_i M_i)^{\dagger}\boldsymbol{c} \\ \text{s.t.} \quad & P\boldsymbol{w} \leq \boldsymbol{d}. \end{aligned} \tag{12}$$

We assume that the optimal design problem is feasible, i.e. there exists a vector $\hat{\boldsymbol{w}} \geq \boldsymbol{0}$ such that $P\hat{\boldsymbol{w}} \leq \boldsymbol{d}$ and $\boldsymbol{c}$ is in the range of $\sum_i \hat{w}_i M_i$. Note that we can assume without loss of generality that $\hat{\boldsymbol{w}} > 0$. Otherwise, this would mean that the constraints $P\boldsymbol{w} \leq \boldsymbol{d},\ \boldsymbol{w} \geq \boldsymbol{0}$ force the equality $w_i = 0$ to hold for some coordinate $i \in [l]$, and in this case we could simply remove the experiment i from the set of available experiments.

We can now express the latter problem as an SDP thanks to the Schur complement lemma:

$$\begin{aligned} \min_{t\in\mathbb{R},\ \boldsymbol{w}\geq\boldsymbol{0}} \quad & t \\ \text{s.t.} \quad & \left(\begin{array}{c|c} \sum_i w_i M_i & \boldsymbol{c} \\ \hline \boldsymbol{c}^T & t \end{array}\right) \succeq 0. \\ & P\boldsymbol{w} \leq \boldsymbol{d}. \end{aligned} \tag{13}$$

Since the optimal t is positive (we exclude the trivial case $\boldsymbol{c} = \boldsymbol{0}$), the latter matrix inequality may be rewritten as

$$\sum_i w_i M_i \succeq \frac{\boldsymbol{c}\boldsymbol{c}^{\boldsymbol{T}}}{t},$$

by using the Schur complement lemma again. Finally, we make the change of variables $\boldsymbol{\mu} = t\boldsymbol{w}$ and Problem (13) is equivalent to

$$\begin{aligned} \min_{\boldsymbol{\mu}\geq\boldsymbol{0},t\geq 0} \quad & t \\ \text{s.t.} \quad & \sum_{i=1}^{l} \mu_i M_i \succeq \boldsymbol{c}\boldsymbol{c}^{\boldsymbol{T}} \\ & P\boldsymbol{\mu} \leq t\boldsymbol{d}. \end{aligned} \tag{14}$$

This problem is exactly in the form of Problem (9), for $C = \boldsymbol{c}\boldsymbol{c}^{\boldsymbol{T}},\ \mu_{l+1} = t,\ \boldsymbol{b} = [0, \ldots, 0, 1]^T \in \mathbb{R}^{l+1},\ M_{l+1} = 0,\ \boldsymbol{h_0} = \boldsymbol{0},\ H = [P, -\boldsymbol{d}]$, and for all $i \in 0, \ldots, l+1,\ R_i = 0$ (we also need to introduce a nonnegative slack variable to handle the inequalities as equalities).

By a similar argument as in the proof of Theorem 2 (see the reduction to Problem (17)), we can assume that $\sum_{i=1}^{l} M_i \succ 0$ (this works because $\boldsymbol{c} \in \text{Range}(\sum_i M_i)$). Therefore, there exists a scalar $\lambda > 0$ such that $\lambda \sum_i M_i \succ \boldsymbol{c}\boldsymbol{c}^T$. We set $\bar{t} = \frac{\lambda}{\min_{i\in[l]} \hat{w}_i}$ ($\bar{t}$ is well defined because $\hat{\boldsymbol{w}} > 0$) and $\overline{\boldsymbol{\mu}} = \bar{t}\hat{\boldsymbol{w}}$, so that $P\overline{\boldsymbol{\mu}} \leq \bar{t}\boldsymbol{d}$, and $\sum_{i=1}^{l} \overline{\mu}_i M_i \succeq \sum_{i=1}^{l} \lambda M_i \succ \boldsymbol{c}\boldsymbol{c}^{\boldsymbol{T}}$. This shows that Problem (14) is strictly feasible. In addition, the corresponding primal problem is clearly feasible (for $\boldsymbol{\lambda} = \boldsymbol{0}$, since $\boldsymbol{b} \geq \boldsymbol{0}$), and thus we can use Corollary 5: the $\boldsymbol{c}-$optimal design problem with resource constraints (12)

reduces to the SOCP (11). We give below this SOCP (with the parameters $\boldsymbol{b}$, M_i and H defined as above), as well as its dual:

$$\begin{aligned} \max_{\boldsymbol{x},\,\boldsymbol{\lambda}} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ & \left\| \begin{bmatrix} 2A_i\boldsymbol{x} \\ \boldsymbol{p_i}^T\boldsymbol{\lambda} - 1 \end{bmatrix} \right\|_2 \leq \boldsymbol{p_i}^T\boldsymbol{\lambda} + 1 \quad (\forall i \in [l]) \\ & \boldsymbol{d}^T\boldsymbol{\lambda} \leq 1, \\ & \boldsymbol{\lambda} \geq \boldsymbol{0}. \end{aligned} \qquad \begin{aligned} \min_{\boldsymbol{\mu}\geq\boldsymbol{0},t\geq 0,\boldsymbol{\alpha}\geq\boldsymbol{0},(\boldsymbol{z_i})_{i\in[l]}} \quad & \sum_{i=1}^{l} \alpha_i + t \\ & \sum_{i=1}^{l} A_i^T\boldsymbol{z_i} = \boldsymbol{c} \\ & P\boldsymbol{\mu} \leq t\boldsymbol{d} \\ & \left\| \begin{bmatrix} \boldsymbol{z_i} \\ \alpha_i - \mu_i \end{bmatrix} \right\|_2 \leq \alpha_i + \mu_i, \end{aligned}$$

where the vectors $\boldsymbol{p_1}, \ldots, \boldsymbol{p_l} \in \mathbb{R}^q$ are the columns of the matrix P. According to the previous change of variable, the optimal design variable $\boldsymbol{w}$ is related to the dual optimal variables $\boldsymbol{\mu}$ and t by the relation $\boldsymbol{w} = t^{-1}\boldsymbol{\mu}$. Moreover, Corollary 5 shows that the optimal value of Problem (12) is the square of the optimal value of these SOCPs.

4 Proofs of the theorems

Proof of Theorem 1. The fact that the problem is feasible if and only if every b_i is nonnegative is clear, since $X = 0$ is always feasible in this case and $M_i \succeq 0, X \succeq 0$, implies $\langle M_i, X\rangle \geq 0$.

Now, we assume that each b_i is nonnegative, and we show that Problem (1) is bounded if and only if $\operatorname{Range} C \subset \operatorname{Range} \sum_i M_i$. The positive semidefiniteness of the matrices M_i implies that there exists matrices A_i ($i \in [l]$) such that $A_i^T A_i = M_i$, and $[A_1^T, \cdots, A_l^T][A_1^T, \cdots, A_l^T]^T = \sum_i M_i$. We also consider a decomposition $C = \sum_{k=1}^r \boldsymbol{c_k}\boldsymbol{c_k}^T$. For any factorization $M = A^T A$ of a positive semidefinite matrix M, it is known that $\operatorname{Range} M = \operatorname{Range} A$, and so the following equivalence relations hold:

$$\begin{aligned} \operatorname{Range} C \subset \operatorname{Range} \sum_i M_i &\Longleftrightarrow \forall k \in [r],\ \boldsymbol{c_k} \in \operatorname{Range}(\sum_i M_i) = \operatorname{Range}([A_1^T, \cdots, A_l^T]) \\ &\Longleftrightarrow \forall k \in [r],\ \boldsymbol{c_k} \in \left(\bigcap_{i=1}^{l} \operatorname{Nullspace}(A_i) \right)^{\perp}. \end{aligned} \tag{15}$$

We first assume that the range inclusion condition does not hold. Relation (15) shows that

$$\exists k \in [r], \exists \boldsymbol{h} \in \mathbb{R}^n : \forall i \in [l], \quad A_i\boldsymbol{h} = 0, \quad \boldsymbol{c_k}^T\boldsymbol{h} \neq 0.$$

Now, notice that $X = \alpha\boldsymbol{h}\boldsymbol{h}^T$ is feasible for all $\alpha > 0$, since $\alpha\langle A_i^T A_i, \boldsymbol{h}\boldsymbol{h}^T\rangle = 0 \leq b_i$. This contradicts the fact that Problem (1) is bounded, because $\langle C, X\rangle \geq \alpha(\boldsymbol{c_k}^T\boldsymbol{h})^2$, and α can be chosen arbitrarily large.

Conversely, if the range inclusion holds, we consider the dual of Problem (1):

$$\begin{aligned} \min_{\boldsymbol{\mu}\geq 0} \quad & \boldsymbol{\mu^T b} \\ \text{s.t.} \quad & \sum_i \mu_i M_i \succeq C. \end{aligned} \tag{16}$$

The range inclusion condition indicates that this problem is feasible, because it implies the existence of a scalar $\lambda > 0$ such that $\lambda \sum_i M_i \succeq C$ (we point out that a convenient value for λ is $\sum_{k=1}^r \boldsymbol{c_k}^T(\sum_i M_i)^\dagger \boldsymbol{c_k}$; this can be seen with the help of the Schur complement lemma). This means that Problem (16) has a finite optimal value $OPT \leq \lambda \sum_i b_i$, and by weak duality, Problem (1) is bounded (its optimal value cannot exceed OPT). ◻

We next prove that if Theorem 4 holds, then so does Theorem 2.

Proof of Theorem 2. We will show that under the conditions of Theorem 2, we can assume without loss of generality that the dual problem (16) is striclty feasible, so that Theorem 4 applies. More precisely, we will see that when the range of C is included in that of $\sum_i M_i$, we can construct an equivalent problem in which the sum of the matrices M_i has full rank, and the conclusion will follow. Let us denote by U a matrix whose columns form an orthonormal basis of the range of $\sum_{i=1}^{l} M_i$ (the matrix U is obtained by taking the eigenvectors corresponding to the nonzero eigenvalues of $\sum_i M_i$). Note that every matrix M_i can be decomposed as $M_i = UM_i'U^T$ for a given matrix M_i', because its range is included in the range of $\sum_i M_i$ (we have $M_i' = U^T M_i U$). The same observation holds for C, which can be decomposed as $C = UC'U^T$ (we have assumed the range inclusion $\mathrm{Range}\, C \subset \mathrm{Range} \sum_i M_i$). Hence, Problem (8) is equivalent to:

$$\begin{aligned} \max_{X \succeq 0} \quad & \langle C', U^T X U \rangle \\ \text{s.t.} \quad & \langle M_i', U^T X U \rangle \leq b_i, \qquad i \in [l], \end{aligned} \tag{17}$$

After the change of variable $Z = U^T X U$ (Z is a positive semidefinite matrix if and only if X is), we obtain a reduced semidefinite packing problem. The projected matrices in the constraints satisfy $\sum_i M_i' = U^T (\sum_i M_i) U \succ 0$. Therefore, there is a real $\lambda > 0$ such that $\lambda \sum_i M_i \succ C$, and the dual of the reduced problem (17) is striclty feasible, while the primal problem is feasible (since every b_i is nonnegative). Hence, Theorem (4) applies, which guarantees the existence of a solution Z of rank at most $r' = \mathrm{rank}\, C' = \mathrm{rank}\, U^T C U \leq \mathrm{rank}\, C = r$ for Problem (17). Finally, $X = UZU^T$ is a solution of the initial problem (1) by construction, and $\mathrm{rank}\, X \leq \mathrm{rank}\, Z \leq r$. □

Before we give the proof of Theorem (4), we need one additional technical lemma, which shows that one can assume without loss of generality that the primal problem is strictly feasible, and that the vector space spanned by the vectors $\boldsymbol{h_0}, \boldsymbol{h_1}, \ldots, \boldsymbol{h_l}$ coincides with the cone generated by the same vectors.

Lemma 6. *We assume that the conditions of Theorem 4 are fulfilled. Then, we can construct a reduced "combined" semidefinite packing problem*

$$\max_{Z \succeq 0,\ Y \succeq 0,\ \boldsymbol{\lambda}} \langle C', Z \rangle + \langle R_0, Y \rangle + \boldsymbol{h_0}^T \boldsymbol{\lambda} \qquad s.t. \qquad \forall i \in \mathcal{I},\ \langle M_i', Z \rangle \leq b_i + \langle R_i, Y \rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda},$$

that has the same optimal value as (8) *and that satisfies the following properties:*

- (i) $\exists (Z' \succ 0, Y' \succ 0, \boldsymbol{\lambda'}) : \ \forall i \in \mathcal{I},\ \langle M_i, Z' \rangle < b_i + \langle R_i, Y' \rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda'}$*;*
- (ii) *The cone K generated by the vectors $(\boldsymbol{h_i})_{i \in \{0\} \cup \mathcal{I}}$ is a vector space.*
- (iii) $\mathrm{rank}\, C' \leq \mathrm{rank}\, C$*;*
- (iv) *There is a matrix U with orthonormal columns such that if (Z, Y) is a solution of the reduced problem, then $(X := UZU^T, Y)$ is a solution of Problem* (8) *(which of course satisfies* $\mathrm{rank}\, X \leq \mathrm{rank}\, Z$*).*

Proof. In this lemma, (i) and (ii) are the properties that we will need to prove Theorem 4. Properties (iii) and (iv) ensure that if the theorem holds for the reduced problem, then the result also holds for the initial problem (8). We handle separately the cases in which the initial problem does not satisfy the property (i) or (ii). If both cases arise simultaneously, we obtain the result of this lemma by applying successively the following two reductions.

Let $(X^*, Y^*, \boldsymbol{\lambda}^*)$ be an optimal solution of Problem (8) ; the existence of a solution is guaranteed by the strict feasibility of the dual problem indeed. We denote by $\mathcal{I}_0 \subset [l]$ the subset of indices for

which $b_i + \langle R_i, Y^* \rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda^*} = 0$ (note that we have $b_i + \langle R_i, Y^* \rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda^*} \geq 0$ for all i because $M_i \succeq 0$ implies $\langle M_i, X^* \rangle \geq 0$). We define $\mathcal{I} := [l] \setminus \mathcal{I}_0$. In Problem (8), we can replace the constraint $\langle M_i, X \rangle \leq b_i + \langle R_i, Y \rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda}$ by $\langle M_i, X \rangle = 0$ for all $i \in \mathcal{I}_0$, since $(X^*, Y^*, \boldsymbol{\lambda^*})$ satisfies this stronger set of constraints. For a feasible positive semidefinite matrix X, this implies $\langle \sum_{i \in \mathcal{I}_0} M_i, X \rangle = 0$, and even $\sum_{i \in \mathcal{I}_0} M_i X = 0$. Therefore, X is of the form UZU^T for some positive semidefinite matrix Z, where the columns of U form an orthonormal basis of the nullspace of $M_0 := \sum_{i \in \mathcal{I}_0} M_i$ (U is obtained by taking the eigenvectors corresponding to the vanishing eigenvalues of M_0). Hence, Problem (8) is equivalent to:

$$\begin{aligned} \max \quad & \langle U^T C U, Z \rangle + \langle R_0, Y \rangle + \boldsymbol{h_0}^T \boldsymbol{\lambda} \\ \text{s.t.} \quad & \langle U^T M_i U, Z \rangle \leq b_i + \langle R_i, Y \rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda}, \qquad i \in \mathcal{I}, \\ & Z \succeq 0, \ Y \succeq 0. \end{aligned} \tag{18}$$

We have thus reduced the problem to one for which $b_i + \langle R_i, Y \rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda} > 0$ for all i, and strict feasibility follows (i.e. property (i) holds, consider $\boldsymbol{\lambda'} = \boldsymbol{\lambda^*}, Y' = Y^* + \eta_1 I$, and $Z' = \eta_2 I$ for sufficiently small reals $\eta_1 > 0$ and $\eta_2 > 0$). Moreover, the projected matrix $C' := U^T C U$ in the objective function has a smaller rank than C (i.e. (iii) holds). Finally, (iv) holds for the reduced problem by construction: if $(Z, Y, \boldsymbol{\lambda})$ is a solution of Problem (18), then $(X = UZU^T, Y, \boldsymbol{\lambda})$ is a solution of Problem (8), both problems have the same optimal value, and of course rank $X \leq$ rank Z.

We now handle the second case, in which Property (ii) does not hold for Problem (8). The set $K = \{ \ [\boldsymbol{h_0}, H]\boldsymbol{v}, \ \boldsymbol{v} \in \mathbb{R}^{l+1}, \boldsymbol{v} \geq \boldsymbol{0}\}$ is a closed convex cone. Hence, it is known that it can be decomposed as $K = L + Q$, where L is a vector space and $Q \subset L^\perp$ is a closed convex pointed cone ($L = K \cap (-K)$ is the *lineality space* of K). The interior of the dual cone Q^* is therefore nonempty, i.e. $\exists \boldsymbol{\lambda} : \forall \boldsymbol{q} \in Q \setminus \{\boldsymbol{0}\}, \boldsymbol{\lambda}^T \boldsymbol{q} > 0$. Let $\boldsymbol{\lambda_0}$ be the orthogonal projection of $\boldsymbol{\lambda}$ on $L^\perp$, so that $\boldsymbol{\lambda_0}^T \boldsymbol{q} = \boldsymbol{\lambda}^T \boldsymbol{q} > 0$ for all $\boldsymbol{q} \in Q \setminus \{\boldsymbol{0}\}$, and $\boldsymbol{\lambda_0}^T \boldsymbol{x} = 0$ for all $\boldsymbol{x} \in L$. Now, we define the set of indices $\mathcal{I} = \{i \in [l] : \boldsymbol{h_i} \in L\}$, and its complement $\mathcal{I}_0 = [l] \setminus \mathcal{I}$. For all $i \in \mathcal{I}_0$, $\boldsymbol{h_i} = \boldsymbol{x_i} + \boldsymbol{q_i}$ for a vector $\boldsymbol{x_i} \in L$ and a vector $\boldsymbol{q_i} \in Q \setminus \{\boldsymbol{0}\}$, so that $\boldsymbol{\lambda_0}^T \boldsymbol{h_i} = \boldsymbol{\lambda_0}^T \boldsymbol{x_i} + \boldsymbol{\lambda_0}^T \boldsymbol{q_i} = \boldsymbol{\lambda_0}^T \boldsymbol{q_i} > 0$. For the indices $i \in \mathcal{I}$, it is clear that $\boldsymbol{\lambda_0}^T \boldsymbol{h_i} = 0$. Finally, since $\boldsymbol{h_0} + H\boldsymbol{\overline{\mu}} = 0$, we have $-\boldsymbol{h_0} \in K$, so that $\boldsymbol{h_0} \in L$ and $\boldsymbol{h_0}^T \boldsymbol{\lambda} = 0$. To sum up, we have proved the existence of a vector $\boldsymbol{\lambda_0}$ for which

$$\forall i \in \{0\} \cup \mathcal{I}, \ \boldsymbol{\lambda_0}^T \boldsymbol{h_i} = 0 \quad \text{and} \quad \forall i \in \mathcal{I}_0, \boldsymbol{\lambda_0}^T \boldsymbol{h_i} > 0.$$

Let $(X^*, Y^*, \boldsymbol{\lambda}^*)$ be an optimal solution of Problem (8). For all positive real t, $(X^*, Y^*, \boldsymbol{\lambda}^* + t\boldsymbol{\lambda_0})$ is also a solution, because it is feasible and has the same objective value. Letting $t \to \infty$, we see that the constraints of the problem that are indexed by $i \in \mathcal{I}_0$ may be removed without changing the optimum. We have thus reduced the problem to one for which (ii) holds.

$\square$

We can now prove the main result of this article (in its extended version for "combined" problems). We will first show that the result holds when each M_i is positive definite, thanks to the complementary slackness relation. Then, the general result is obtained by continuity. We point out at the end of this section the sketch of an alternative proof of Theorem 2 for the case in which $r = 1$, based on the bidual of Problem (1) and Schur complements, that shows directly that Problem (1) reduces to the SOCP (2).

Proof of Theorem 4. By Lemma 6, we may assume without loss of generality that $K = \text{cone}\{\boldsymbol{h_0}, \dots, \boldsymbol{h_l}\} \supset -K$ and that the primal problem is strictly feasible. The strict feasibility of both the primal and the dual problem ensures the existence of solutions for Problems (8) and (9), and strong duality holds because the Slater condition is fulfilled. This means that there is no duality gap between (8) and (9), and that the pairs of primal and dual solutions $\big((X^*, Y^*, \boldsymbol{\lambda^*}), \boldsymbol{\mu}^*\big)$ are

characterized by the Karush-Kuhn-Tucker (KKT) conditions:

$$\begin{aligned}
\text{Primal Feasibility:}\quad & \forall i \in [l], \quad \langle M_i, X^* \rangle \leq b_i + \langle R_i, Y^* \rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda^*}, \\
& X^* \succeq 0,\ Y^* \succeq 0; \\
\text{Dual Feasibility:}\quad & \boldsymbol{\mu^*} \geq 0, \quad \sum_{i=1}^{l} \mu_i^* M_i \succeq C, \quad R_0 + \sum_{i=1}^{l} \mu_i^* R_i \preceq 0, \quad \boldsymbol{h_0} + H\boldsymbol{\mu^*} = 0; \\
\text{Complementary Slackness:}\quad & (\sum_{i=1}^{l} \mu_i^* M_i - C)\ X^* = 0, \qquad (R_0 + \sum_{i=1}^{l} \mu_i^* R_i)\ Y^* = 0.
\end{aligned}$$

Now, we consider the case in which $M_i \succ 0$ for all i, and we choose an arbitrary pair of primal and dual optimal solutions $\big((X^*, Y^*, \boldsymbol{\lambda^*}), \boldsymbol{\mu}^*\big)$. The dual feasibility relation implies $\boldsymbol{\mu}^* \neq \mathbf{0}$, and so $\sum_i \mu_i^* M_i$ is a positive definite matrix (we exclude the trivial case $C = 0$). Since C is of rank r, we deduce that

$$\operatorname{rank}(\sum_i \mu_i^* M_i - C) \geq n - r.$$

Finally, the complementary slackness relation indicates that the columns of X^* belong to the nullspace of $(\sum_i \mu_i^* M_i - C)$, which is a vector space of dimension at most $n - (n - r) = r$, and so we conclude that $\operatorname{rank} X^* \leq r$.

Note that the same reasoning as above shows that if $\min_{i\in[l]}\ \operatorname{rank}\ M_i = \overline{r}$, then we have $\operatorname{rank}(\sum_i \mu_i^* M_i - C) \geq \overline{r} - r$, and every solution of Problem (8) must be of rank at most $n - \overline{r} + r$.

We now turn to the study of the general case in which $M_i \succeq 0$. To this end, we consider the perturbed problems

$$\begin{aligned}
\max\quad & \langle C, X \rangle + \langle R_0, Y \rangle + \boldsymbol{h_0}^T \boldsymbol{\lambda} \\
\text{s.t.}\quad & \langle M_i + \varepsilon I, X \rangle \leq b_i + \langle R_i, Y \rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda} \qquad i \in [l], \\
& X \succeq 0,\ Y \succeq 0,
\end{aligned} \tag{P_ε}$$

and

$$\begin{aligned}
\min_{\boldsymbol{\mu}\geq 0}\quad & \sum_{i=1}^{l} \mu_i b_i, \\
\text{s.t.}\quad & \sum_{i=1}^{l} \mu_i (M_i + \varepsilon I) \succeq C,, \\
& R_0 + \sum_{i=1}^{l} \mu_i R_i \preceq 0, \\
& \boldsymbol{h_0} + H\boldsymbol{\mu} = \mathbf{0}.
\end{aligned} \tag{D_ε}$$

where $\varepsilon \geq 0$. Note that the strict feasibility of the unperturbed problems (8) and (9) implies that of (P_ε) and (D_ε) on a neighborhood $\varepsilon \in [0, \varepsilon_0]$, $\varepsilon_0 > 0$. We denote by $\big((X^\varepsilon, Y^\varepsilon, \boldsymbol{\lambda^\varepsilon}), \boldsymbol{\mu^\varepsilon}\big)$ a pair of primal and dual solutions of (P_ε)–(D_ε). If $\varepsilon > 0$, $M_i + \varepsilon I \succ 0$ and it follows from the previous discussion that X^ε is of rank at most r. We show below that we can choose the optimal variables $(X^\varepsilon, Y^\varepsilon, \boldsymbol{\lambda^\varepsilon}, \boldsymbol{\mu^\varepsilon})_{\varepsilon\in]0,\varepsilon_0]}$ within a bounded region, so that we can construct a converging subsequence $(X^{\varepsilon_k}, Y^{\varepsilon_k}, \boldsymbol{\lambda^{\varepsilon_k}}, \boldsymbol{\mu^{\varepsilon_k}})_{k\in\mathbb{N}}$, $\varepsilon_k \to 0$ from these variables. To conclude, we will see that the limit $(X_0, Y_0, \boldsymbol{\lambda^0}, \boldsymbol{\mu^0})$ satisfies the KKT conditions for Problems (8)–(9), and that X_0 is of rank at most r.

Let us denote the optimal value of Problems (P_ε)–(D_ε) by $OPT(\varepsilon)$. Since the constraints of the primal problem becomes tighter when ε grows, it is clear that $OPT(\varepsilon)$ is non-increasing with ε, so that

$$\forall \varepsilon \in [0,\varepsilon_0],\ OPT(\varepsilon_0) \leq OPT(\varepsilon) \leq OPT(0).$$

Now let $\varepsilon \in]0,\varepsilon_0]$. By dual strict feasibility assumption (i), there exists a vector $\overline{\boldsymbol{\mu}} \geq \mathbf{0}$ such that

$$\sum_i \overline{\mu}_i (M_i + \varepsilon I) \succeq \sum_i \overline{\mu}_i M_i \succ C, \quad \text{and} \quad R_0 + \sum_i \overline{\mu}_i R_0 \prec 0. \tag{19}$$

Therefore, we have

$$\begin{aligned} OPT(\varepsilon) = \langle C, X^\varepsilon\rangle + \langle R_0, Y^\varepsilon\rangle + \boldsymbol{h_0}^T \boldsymbol{\lambda^\varepsilon} &\leq \big\langle \sum_i \overline{\mu}_i (M_i + \varepsilon I), X^\varepsilon \big\rangle + \langle R_0, Y^\varepsilon\rangle + \boldsymbol{h_0}^T \boldsymbol{\lambda^\varepsilon} \\ &\leq \sum_i \overline{\mu}_i \big(b_i + \langle R_i, Y^\varepsilon\rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda^\varepsilon}\big) + \langle R_0, Y^\varepsilon\rangle + \boldsymbol{h_0}^T \boldsymbol{\lambda^\varepsilon} \\ &= \overline{\boldsymbol{\mu}}^T \boldsymbol{b} + \langle \sum_i \overline{\mu}_i R_i + R_0, Y^\varepsilon\rangle + (\underbrace{\boldsymbol{h_0} + H\overline{\boldsymbol{\mu}}}_{=\mathbf{0}})^T \boldsymbol{\lambda^\varepsilon}, \end{aligned}$$

where the first inequality follows from (19), and the second one from the feasibility condition $\langle M_i + \varepsilon I, X^\varepsilon\rangle \leq b_i + \langle R_i, Y^\varepsilon\rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda^\varepsilon}$. The assumption (19) moreover implies that $-(\sum_i \overline{\mu}_i R_i + R_0)$ is positive definite, so that its smallest eigenvalue λ is positive, and

$$\lambda \text{ trace } Y^\varepsilon \leq \big\langle -(\sum_i \overline{\mu}_i R_i + R_0), Y^\varepsilon \big\rangle \leq \overline{\boldsymbol{\mu}}^T \boldsymbol{b} - OPT(\varepsilon) \leq \overline{\boldsymbol{\mu}}^T \boldsymbol{b} - OPT(\varepsilon_0).$$

This shows that the trace of Y^ε is bounded, and so $Y^\varepsilon \succeq 0$ is bounded.

Similarly, to bound X^ε, we write:

$$\begin{aligned} \langle \sum_i \overline{\mu}_i M_i - C, X^\varepsilon \rangle &\leq \langle \sum_i \overline{\mu}_i (M_i + \varepsilon I) - C, X^\varepsilon \rangle \\ &= \langle \sum_i \overline{\mu}_i (M_i + \varepsilon I), X^\varepsilon \rangle - OPT(\varepsilon) + \langle R_0, Y^\varepsilon\rangle + \boldsymbol{h_0}^T \boldsymbol{\lambda^\varepsilon} \\ &\leq \sum_i \overline{\mu}_i \big(b_i + \langle R_i, Y^\varepsilon\rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda^\varepsilon}\big) - OPT(\varepsilon) + \langle R_0, Y^\varepsilon\rangle + \boldsymbol{h_0}^T \boldsymbol{\lambda^\varepsilon} \\ &= \overline{\boldsymbol{\mu}}^T \boldsymbol{b} - OPT(\varepsilon) + \underbrace{\langle \sum_i \overline{\mu}_i R_i + R_0, Y^\varepsilon \rangle}_{\leq 0} + (\underbrace{\boldsymbol{h_0} + H\overline{\boldsymbol{\mu}}}_{=\mathbf{0}})^T \boldsymbol{\lambda^\varepsilon}, \end{aligned}$$

where the first equality comes from the expression of $OPT(\varepsilon)$. The matrix $\sum_i \overline{\mu}_i M_i - C$ is positive definite and its smallest eigenvalue λ' is therefore positive. Hence,

$$\lambda' \text{ trace } X^\varepsilon \leq \overline{\boldsymbol{\mu}}^T \boldsymbol{b} - OPT(\varepsilon) \leq \overline{\boldsymbol{\mu}}^T \boldsymbol{b} - OPT(\varepsilon_0),$$

and this shows that the matrix $X^\varepsilon \succeq 0$ is bounded.

Now, note that the feasibility of $\boldsymbol{\lambda^\varepsilon}$ implies that the quantity $b_i + \langle R_i, Y^\varepsilon\rangle + \boldsymbol{h_i}^T \boldsymbol{\lambda^\varepsilon}$ is nonnegative for all $i \in [l]$. Since Y^ε is bounded, we deduce the existence of a lower bound $m_i \in \mathbb{R}$ such that $\boldsymbol{h_i}^T \boldsymbol{\lambda^\varepsilon} \geq m_i$ ($\forall i \in [l]$). Similarly, since $\boldsymbol{h_0}^T \boldsymbol{\lambda^\varepsilon} \geq OPT(\varepsilon_0) - \langle C, X^\varepsilon\rangle - \langle R_0, Y^\varepsilon\rangle$, there is a scalar m_0 such that $\boldsymbol{h_0}^T \boldsymbol{\lambda^\varepsilon} \geq m_0$. We now use the fact that every vector $(-\boldsymbol{h_i})$ may be written as a positive combination of the $\boldsymbol{h_k}$, ($k \in \{0\} \cup [l]$), and we obtain that the quantities $\boldsymbol{h_i}^T \boldsymbol{\lambda^\varepsilon}$ are also bounded from above. Let us denote by H_0 the matrix $[\boldsymbol{h_0}, H]$; we have just proved that the vector $H_0^T \boldsymbol{\lambda^\varepsilon}$ is bounded:

$$\exists \overline{m} \in \mathbb{R}: \quad \|H_0^T \boldsymbol{\lambda^\varepsilon}\|_2 \leq \overline{m}$$

(the latter bound does not depend on ε). Note that one may assume without loss of generality that $\boldsymbol{\lambda^\varepsilon} \in \operatorname{Range} H_0$ (otherwise we consider the projection $\boldsymbol{\lambda_P^\varepsilon}$ of $\boldsymbol{\lambda^\varepsilon}$ on $\operatorname{Range} H_0$ which is also a solution since $H_0^T\boldsymbol{\lambda^\varepsilon} = H_0^T\boldsymbol{\lambda_P^\varepsilon}$. We know from the Courant-Fisher theorem that the smallest positive eigenvalue of $H_0H_0^T$ satisfies:

$$\lambda_{\min}^{>}(H_0H_0^T) = \min_{\boldsymbol{v}\in \operatorname{Range} H_0\setminus\{\boldsymbol{0}\}} \frac{\boldsymbol{v}^T H_0 H_0^T \boldsymbol{v}}{\boldsymbol{v}^T\boldsymbol{v}}.$$

Therefore, since we have assumed $\boldsymbol{\lambda^\varepsilon} \in \operatorname{Range} H_0$:

$$\|\boldsymbol{\lambda^\varepsilon}\|^2 \leq \frac{\|H_0^T\boldsymbol{\lambda^\varepsilon}\|^2}{\lambda_{\min}^{>}(H_0H_0^T)} \leq \frac{\overline{m}^2}{\lambda_{\min}^{>}(H_0H_0^T)}.$$

It remains to show that the dual optimal variables $\boldsymbol{\mu^\varepsilon}$ are bounded. Lemma (6) ensures the existence of a matrix $\overline{Y} \succeq 0$ and a vector $\overline{\lambda}$ such that

$$\forall i \in [l],\ \langle R_i, \overline{Y}\rangle + b_i + \boldsymbol{h_i}^T\overline{\boldsymbol{\lambda}} = \eta_i > 0.$$

By dual feasibility, $R_0 + \sum_i \mu_i^\varepsilon R_i$ is a negative semidefinite matrix, and we have:

$$0 \geq \langle R_0, \overline{Y}\rangle + \sum_{i=1}^{l} \mu_i^\varepsilon \langle R_i, \overline{Y}\rangle = \langle R_0, \overline{Y}\rangle + \sum_{i=1}^{l} \mu_i^\varepsilon(\eta_i - b_i - \boldsymbol{h_i}^T\overline{\boldsymbol{\lambda}}).$$

Hence, we have the following inequalities:

$$\begin{aligned}\forall k \in [l],\ \eta_k\mu_k^\varepsilon \leq \sum_{i=1}^{l}\eta_i\mu_i^\varepsilon &\leq \boldsymbol{b}^T\boldsymbol{\mu^\varepsilon} + \overline{\boldsymbol{\lambda}}^T H\boldsymbol{\mu^\varepsilon} - \langle R_0, \overline{Y}\rangle \\ &= OPT(\varepsilon) - \overline{\boldsymbol{\lambda}}^T\boldsymbol{h_0} - \langle R_0, \overline{Y}\rangle \\ &\leq OPT(0) - \overline{\boldsymbol{\lambda}}^T\boldsymbol{h_0} - \langle R_0, \overline{Y}\rangle,\end{aligned}$$

and we have shown that $\boldsymbol{\mu^\varepsilon} \geq \boldsymbol{0}$ is bounded.

We can therefore construct a sequence of pairs of primal and dual optimal solutions $(X^{\varepsilon_k}, Y^{\varepsilon_k}, \boldsymbol{\lambda^{\varepsilon_k}}, \boldsymbol{\mu^{\varepsilon_k}})_{k\in\mathbb{N}}$ that converges, with $\varepsilon_k \underset{k\to\infty}{\longrightarrow} 0$, $\varepsilon_k > 0$. The limit X_0 of this sequence is of rank at most r, because the rank is a lower semicontinuous function and $\operatorname{rank} X^{\varepsilon_k} \leq r$ for all $k \in \mathbb{N}$. It remains to show that X_0 is a solution of Problem (8). The $\varepsilon-$perturbed KKT conditions must hold for all $k \in \mathbb{N}$, and so they hold for the pair $\big((X_0, Y_0), \boldsymbol{\lambda^0}, \boldsymbol{\mu^0}\big)$ by taking the limit (the limit of positive semidefinite matrices is a positive semidefinite matrix because $\mathbb{S}_n^+$ is closed). This concludes the proof. $\square$

Sketch of an alternative proof of Theorem 2 when $r = 1$

When $r = 1$, there is a vector $\boldsymbol{c}$ such that $C = \boldsymbol{cc^T}$ and the dual problem of (1) takes the form:

$$\begin{aligned}\min_{\boldsymbol{\mu}\geq 0}\quad & \boldsymbol{\mu^T b} \\ \text{s.t.}\quad & \boldsymbol{cc^T} \preceq \sum_i \mu_i M_i.\end{aligned} \tag{20}$$

Now, setting $t = \boldsymbol{\mu^T b}$, and $\boldsymbol{w} = \frac{\boldsymbol{\mu}}{t}$, so that the new variable $\boldsymbol{w}$ satisfies $\boldsymbol{w^T b} = 1$, the constraint of the previous problem becomes $\frac{\boldsymbol{cc^t}}{t} \preceq \sum_i w_i M_i$. This matrix inequality, together with the fact that

the optimal t is positive, can be reformulated thanks to the Schur complement lemma, and (20) is equivalent to:

$$\begin{aligned} \min_{t\in\mathbb{R},\boldsymbol{w}\geq\mathbf{0}} \quad & t \\ \text{s.t.} \quad & \left(\begin{array}{c|c} \sum_i w_i M_i & \boldsymbol{c} \\ \hline \boldsymbol{c^T} & t \end{array}\right) \succeq 0. \\ & \boldsymbol{w^T b} = 1. \end{aligned} \tag{21}$$

We dualize this SDP once again to obtain the bidual of Program (1) (strong duality holds):

$$\begin{aligned} \max_{\beta\in\mathbb{R},Z\in\mathbb{S}_{n+1}^+} \quad & -\beta - 2\boldsymbol{v^T c} \\ \text{s.t.} \quad & \langle W, M_i\rangle \leq \beta b_i, \quad i\in[l] \\ & Z = \left(\begin{array}{c|c} W & \boldsymbol{v} \\ \hline \boldsymbol{v^T} & 1 \end{array}\right) \succeq 0. \end{aligned} \tag{22}$$

We notice that the last matrix inequality is equivalent to $W \succeq \boldsymbol{vv^T}$, using a Schur complement. Since $M_i \succeq 0$, we can assume that $W = \boldsymbol{vv^T}$ without loss of generality, and (22) becomes:

$$\begin{aligned} \max_{\beta\in\mathbb{R},\boldsymbol{v}\in\mathbb{R}^n} \quad & -\beta - 2\boldsymbol{v^T c} \\ \text{s.t.} \quad & \|A_i\boldsymbol{v}\|^2 \leq \beta b_i, \quad i = 1 \in [l], \end{aligned} \tag{23}$$

where A_i is a matrix such that $A_i^T A_i = M_i$.

We now define the new variables $\alpha = \sqrt{\beta}$, and $\boldsymbol{x} = \frac{\boldsymbol{v}}{\alpha}$, so that (23) becomes:

$$\begin{aligned} \max_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \left(\max_{\alpha} -\alpha^2 - 2\alpha\boldsymbol{x^T c}\right) \\ \text{s.t.} \quad & \|A_i\boldsymbol{x}\| \leq \sqrt{b_i}, \quad i = 1 \in [l]. \end{aligned} \tag{24}$$

The reader can finally verify that the value of the max within parenthesis is $(\boldsymbol{c^T x})^2$, and we have proved that SDP (1) reduces to SOCP (2). By the way, this guarantees that SDP (1) has a rank-one solution. □

Acknowledgment

The author thanks Stéphane Gaubert for his useful comments and enlightening discussions, as well as for his warm support. He also expresses his gratitude to two anonymous referees. In a previous version, the main result was restricted to the case in which $r = 1$. One referee suggested an alternative proof with an elegant complementary slackness argument, which led to the more general statement of Theorem 2. The author also thanks the second referee for his useful remarks, which led to clarify the consequences of these results.

References

[1] S. Arora, S. Rao, and U. Vazirani. Expander ows, geometric embeddings, and graph partitionings. In *Journal of the ACM*, 56(2):article No. 5, 37 pages, 2009.

[2] A. d'Aspremont, L. El Ghaoui, M.I. Jordan, and G.R.G. Lanckriet. A direct formulation for sparse PCA using semidenite programming, *SIAM Review*, 49(3):434–448, 2007.

[3] A. I. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete and computational Geometry*, 13:189–202, 1995.

[4] M. Bouhtou, S. Gaubert, and G. Sagnol. Optimization of network traffic measurement: a semidefinite programming approach. In *Proceedings of the International Conference on Engineering Optimization (ENGOPT 2008)*, Rio De Janeiro, Brazil. ISBN 978-85-7650-152-7.

[5] S. Burer and R.D.C. Monteiro. A Nonlinear Programming Algorithm for Solving Semidefinite Programs Via Low-Rank Factorization. *Mathematical Programming (series B)*, 95(2):329-357, 2003.

[6] S. Burer and R.D.C. Monteiro. Local Minima and Convergence in Low-Rank Semidefinite Programming. *Mathematical Programming (series A)*, 103(3):427-444, 2005.

[7] M.X. Goemans and S.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problem using semidefinite programming. *Journal of the ACM*, 42:1115–1145, 1995.

[8] G. Iyengar, D. J. Phillips, and C. Stein. Approximation algorithms for semidefinite packing problems with applications to maxcut and graph coloring. In *Proceedings of the 11th conference on Integer Programming and Combinatorial Optimization*, volume 3509 of *Lecture Notes in Computer Science*, pages 152–166. Springer, 2005.

[9] D. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidenite programming. *Journal of the ACM*, 45(2):246265, 1998.

[10] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25:1–7, 1979.

[11] A. Nemirovski, C. Roos, and T. Terlaky. On maximization of quadratic form over intersection of ellipsoids with common center. *Mathematical programming*, 86:463–473, 1999.

[12] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Mathematics of Operations Research*, (23):339–358, 1998.

[13] F. Pukelsheim. On Linear Regression Designs which maximize information. *Journal of statistical planning and inferrence*,4:339–364,1980.

[14] F. Pukelsheim. *Optimal Design of Experiments*. Wiley, 1993.

[15] P. Richtarik. Simultaneously solving seven optimization problems in relative scale. Optimization online, 2185, 2008.

[16] G. Sagnol Computing Optimal Designs of multiresponse Experiments reduces to Second-Order Cone Programming, Submitted, arXiv.org:0912.5467, 2009.

[17] G. Sagnol, S. Gaubert, and M. Bouhtou. Optimal monitoring on large networks by Successive c-Optimal Designs. To appear in the *proceedings of the 22nd international teletraffic congress (ITC22)*, Amsterdam, The Netherlands, September 2010. Preprint: http://www.cmap.polytechnique.fr/~sagnol/papers/ITC22_submitted.pdf

[18] J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999.

[19] M. Szegedy, A note on the ϑ number of Lovász and the generalized Delsarte bound. In *SFCS 94: Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, Washington, DC, USA, IEEE Computer Society, pp. 3639, 1994.

[20] L. Vandenberghe, S. Boyd, and S. Wu. Determinant maximization with linear matrix inequality constraints. *SIAM Journal on Matrix Analysis and Applications*, 19:499–533, 1998.